

MORA: an Energy-Aware Slack Reclamation Scheme for Scheduling Sporadic Real-Time Tasks upon Multiprocessor Platforms

Vincent Nelis^{1,2}

Joël Goossens¹

Abstract

In this paper, we address the global and preemptive energy-aware scheduling problem of sporadic constrained-deadline tasks on DVFS-identical multiprocessor platforms. We propose an online slack reclamation scheme which profits from the discrepancy between the worst- and actual-case execution time of the tasks by slowing down the speed of the processors in order to save energy. Our algorithm called *MORA* takes into account the application-specific consumption profile of the tasks. We demonstrate that *MORA* does not jeopardize the system schedulability and we show by performing simulations that it can save up to 32% of energy (in average) compared to execution without using any energy-aware algorithm.

1. Introduction

Context of the study. Nowadays, many modern processors can operate at various supply voltages, where different supply voltages lead to different clock frequencies and to different processing speeds. Since the power consumption of a processor is usually a convex and increasing function of its speed, the slower its speed is, the less its consumption is [17]. Among the most recent and popular such processors, one can cite the Intel PXA27x processor family [21], used by many PDA devices [20].

Many computer systems, especially embedded systems, are now equipped with such voltage (speed) scaling processors and adopt various energy-efficient strategies for managing their applications intelligently. Moreover, many recent energy-constrained embedded systems are built upon multiprocessor platforms because of their high-computational requirements. As pointed out in [10, 11], another advantage is that multiprocessor systems are

more energy efficient than equally powerful uniprocessor platforms, because raising the frequency of a single processor results in a *multiplicative* increase of the consumption while adding processors leads to an *additive* increase.

Supported by this emerging technology, the Dynamic Voltage and Frequency Scaling (DVFS) [15] framework becomes a major concern for multiprocessor power-aware embedded systems. For real-time systems, this framework consists in reducing the system energy consumption by adjusting the working voltage and frequency of the processors, while respecting all the timing constraints.

Previous work. There are a large number of researches about the uniprocessor energy-aware real-time scheduling problem [5, 9, 22, 23, 32]. Among those, many *slack reclamation* approaches have been developed over the years. Such techniques dynamically collect the unused computation times at the end of each early task completion and share it among the remaining pending tasks. Examples of such approaches include the ones proposed in [5, 27, 28, 33]. Some reclaiming algorithms even anticipate the early completion of tasks for further reducing the CPU speed [5, 27], some having different levels of “aggressiveness” [5].

In [15], Kuo et al. propose a state-of-art about energy-aware algorithms in *multiprocessor* environment. As it is mentioned in this state-of-art, many studies (see for instance [14, 16, 17, 18, 34, 31]) consider the *frame-based task model*, i.e., all the tasks share a common deadline and this “frame” is indefinitely repeated. Among the most interesting studies which consider this task model, Zhu et al. [34] explored *online* slack reclamation schemes (i.e., running during the system execution) for dependent and independent tasks. In [18], Kuo et al. propose a set of energy-efficient scheduling algorithms with different task remapping and slack reclamation schemes. In [17], the authors address independent tasks, where task migrations are not allowed. In [14], the authors provide some techniques with and without allowing task migration, while assuming that tasks share the same power consumption function and each processor may run at a selected speed, independently from the speeds of the others. In [16], the authors con-

¹Université Libre de Bruxelles (U.L.B.) CP 212, Department of Computer Science, 50 Av. F.D. Roosevelt, B-1050 Brussels, Belgium.

²Supported by the Belgian National Science Foundation (F.N.R.S.) under a F.R.I.A. grant.

sider that tasks are allowed to have different power consumption functions. In [31], energy-aware multiprocessor scheduling of frame-based tasks was explored for multiprocessor architectures, in which all the processors must share the same speed at any time. Finally the authors of [13] propose a slack reclamation scheme for identical multiprocessor platforms, while considering frame-based tasks of which the distribution of the computation times is assumed to be known.

Targeting a *sporadic task model*, Anderson and Baruah [3] explored the trade-off between the total energy consumption of task executions and the number of required processors, where all the tasks run at the same common speed. In previous work [26], we provided a technique that determines the minimum common offline speed for every task under global-EDF policy [7], while considering identical multiprocessor platforms. Furthermore, we proposed in the same study an *online* algorithm called MOTE which was, to the best of our knowledge, the first to address the global and preemptive energy-aware scheduling problem of sporadic constrained-deadlines tasks on multiprocessors. The main idea of MOTE is to anticipate at run-time the coming idle instants in the schedule in order to reduce the processors speed accordingly. This algorithm cannot be considered as a slack reclamation scheme since it does not directly take advantage from early tasks completion, but it can be combined with slack reclaiming techniques (and in particular with MORA) in order to improve the energy savings.

Contribution of the paper. In this paper, we propose a *slack reclamation scheme* called MORA for the global and preemptive energy-aware scheduling problem of *sporadic constrained-deadline real-time tasks* on a *fixed number of DVFS-capable processors*. According to [15] and to the best of our knowledge, this is the first work which addresses a slack reclamation scheme in this context. Although most previous studies on multiprocessor energy-efficient scheduling assumed that the actual execution time of a task is equal to its Worst-Case Execution Time (WCET), such that those in [2, 6, 14, 31] for instance, this work is motivated by the scheduling of tasks in practice, where tasks might usually complete earlier than their WCET [5, 34]. The proposed algorithm MORA is an *online* scheme which exploits early task completions by using as much as possible the unused time to reduce the speed of the processors. Although it has been inspired from the *uniprocessor* “Dynamic Reclaiming Algorithm” (DRA) proposed in [5], the way in which it profits from the unused time is very different from the DRA since MORA takes into account the application-specific consumption profile of the tasks.

Organization of the paper. The document is organized as follows: in Section 2, we introduce our model of computation, in particular our task and platform model; in Sec-

Processor Type	Intel XScale [1]				
Frequency (MHz): f_k	150	400	600	800	1000
Speed: s_k	0.15	0.4	0.6	0.8	1.0
Voltage (V)	0.75	1.0	1.3	1.6	1.8
Power in run mode (mW): $P(s_k)$	80	170	400	900	1600
Power in idle mode (mW): P_{idle}	40				

Table 1. Intel XScale characteristics

tion 3, we present our online slack reclamation technique called MORA and we prove its correctness; in Section 4, we present our simulation results and in Section 5, we introduce future research directions and we conclude.

2. Model of computation

2.1. Platform model

We consider multiprocessor platforms composed of a known and fixed number m of DVFS-*identical* processors $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_m\}$. “DVFS-identical” means that (i) all the processors have the same profile (in term of consumption, computational capabilities, etc.) and are interchangeable, (ii) two processors running at a same frequency execute the same amount of execution units, and (iii) all the processors have the same minimal and maximal operating frequency denoted by f_{\min} and f_{\max} , respectively. The processors are referred to as *independent*, with the interpretation that they can operate at different frequencies at the same time [29, 24]. Furthermore, we assume that each processor can dynamically adapt its operating frequency (and voltage) at any time during the system execution, independently from each other. The time overheads on frequency (voltage) switching are assumed to be negligible, such as in many researches [4, 9, 25, 32, 35].

We define the notion of *speed* s of a processor as the ratio of its operating frequency f over its maximal frequency, i.e.: $s \stackrel{\text{def}}{=} \frac{f}{f_{\max}}$ – with the interpretation that a job that executes on a processor running at speed s for R time units completes $s \times R$ execution units. When only K discrete frequencies are available to a processor, they are sorted in the increasing order of frequency and denoted by f_1, \dots, f_K . For each frequency f_k such that $1 \leq k \leq K$, we denote by s_k the corresponding speed (i.e., $s_k \stackrel{\text{def}}{=} \frac{f_k}{f_{\max}}$) and by $P(s_k)$ the power consumption (energy consumption rate) per second while the processor is running at speed s_k . The available frequencies and the corresponding core voltages of the Intel XScale processor [1] that will be used in our experiments are outlined in Table 1. Notice that, from our definition of the processor speed, s_{\max} is $\frac{f_{\max}}{f_{\max}} = 1$ whatever the considered processor. Moreover, due to the finite

number of speeds that are available to any practical processor, any speed s computed by any energy-aware algorithm must be translated into one of the available speeds. In this work, this translation is performed by the function $\hat{S}(s) \stackrel{\text{def}}{=} \min\{s_i \mid s_i \geq s\}$.

2.2. Application model

A real-time system τ is a set of n functionalities denoted by $\{\tau_1, \tau_2, \dots, \tau_n\}$. Every functionality τ_i is modeled by a *sporadic constrained-deadline* task characterized by three parameters (C_i, D_i, T_i) – a Worst-Case Execution Time (WCET) C_i at maximal processors speed s_{\max} (expressed in milliseconds for instance), a minimal inter-arrival delay T_i and a relative deadline $D_i \leq T_i$ – with the interpretation that the task τ_i generates successive *jobs* $\tau_{i,j}$ (with $j = 1, \dots, \infty$) arriving at times $a_{i,j}$ such that $a_{i,j} \geq a_{i,j-1} + T_i$ (with $a_{i,1} \geq 0$), each such job has a worst-case execution time of at most C_i time units (at maximal processors speed s_{\max}), and must be completed at (or before) its absolute deadline noted $D_{i,j} \stackrel{\text{def}}{=} a_{i,j} + D_i$. According to our definition of the processors speed, a processor running at speed $s_{\max} = 1$ may take up to C_i time units to complete a job $\tau_{i,j}$ and, at a given speed s , its WCET is $\frac{C_i}{s}$. Notice that, since $D_i \leq T_i$, successive jobs of any task τ_i do not interfere with each other.

We define the *density* δ_i of the task τ_i as the ratio of its WCET at maximal speed s_{\max} over its deadline, i.e., $\delta_i \stackrel{\text{def}}{=} \frac{C_i}{D_i}$. We assume that this ratio is not larger than 1 for every task, since a task with a density larger than 1 is never able to meet its deadlines (since task parallelism is forbidden in this work). The *maximal density* $\delta_{\max}(\tau)$ of the system is defined as $\delta_{\max}(\tau) \stackrel{\text{def}}{=} \max_{i=1}^n \{\delta_i\}$ and its *total density* is defined as $\delta_{\text{sum}}(\tau) \stackrel{\text{def}}{=} \sum_{i=1}^n \delta_i$. In our study, all the tasks are assumed to be *independent*, i.e., there is no communication, no precedence constraint and no shared resource (except the processors) between them.

At any time t in any schedule \mathcal{S} , a job $\tau_{i,j}$ is said to be *active* iff $a_{i,j} \leq t$ and it is not completed yet in \mathcal{S} . Moreover, an active job is said to be *running* at time t in \mathcal{S} if it is executing on a processor. Otherwise, the active job is pending in a ready-queue of the operating system and we say that it is *waiting*. Furthermore, a job is said to be *dispatched* at time t in \mathcal{S} if it passes from the waiting state to the running state at time t .

Although certain benchmarks provide measured power consumption, we should not ignore that different applications may have different instruction sequences and require different function units in the processor, thus leading to different dynamic consumption profiles. As it was already done in [30], we hence introduce a measurable parameter e_i for each task τ_i that reflects this application-specific

power difference between the applications and the measured benchmark. Accordingly, the consumption of any task τ_i executed for 1 time unit at speed s_k can be estimated by $e_i \cdot (P(s_k) - P_{\text{idle}}) + P_{\text{idle}}$ [30], where $P(s)$ and P_{idle} are defined as in Table 1. In the remainder of this paper, we denote by $E_i(R, s_k)$ the energy consumed by the task τ_i when executed for R time units at speed s_k and we define it as $E_i(R, s_k) \stackrel{\text{def}}{=} R \cdot (e_i \cdot (P(s_k) - P_{\text{idle}}) + P_{\text{idle}})$. As we will see in Section 3.3, MORA uses these energy consumption functions in order to improve the energy saving that it provides. This improvement makes MORA very different from the *uniprocessor* dynamic reclaiming algorithm DRA proposed in [5].

2.3. Scheduling specifications

We consider in this study the *global* scheduling problem of sporadic constrained-deadlines tasks on multiprocessor platforms. “Global” scheduling algorithms, on the contrary to partitioned algorithms, allow different tasks and *different* jobs of the same task to be executed upon *different* processors. Furthermore, we consider *preemptive* scheduling and *Fixed Job-level Priority* assignment (FJP), with the following interpretations. In the *preemptive global scheduling problem*, every job can start its execution on any processor and *may migrate at run-time* to any other processor if it gets meanwhile preempted by a higher-priority job. We assume in this paper that preemptions are carried out with no loss or penalty. *Fixed Job-level Priority assignment* means that the scheduler assigns a priority to jobs as soon as they arrive and every job keeps its priority constant until it completes. *Global Deadline Monotonic* and *Global Earliest Deadline First* [7] are just some examples of such scheduling algorithms.

3. The Multiprocessor Online Reclaiming Algorithm (MORA)

3.1. Notations

During the system execution, every *active* job $\tau_{i,j}$ has two associated speeds noted $s_{i,j}$ and $s_{i,j}^{\text{off}}$. The speed $s_{i,j}$ denotes *the speed that a processor adopts while executing $\tau_{i,j}$* . We assume that these execution speeds $s_{i,j}$ can be modified at any time during the system execution, even during the execution of $\tau_{i,j}$, and it is instantaneously reflected on the processor speed. On the other hand, the speed $s_{i,j}^{\text{off}}$ is the *offline precomputed execution speed* of $\tau_{i,j}$, in the sense that the value of $s_{i,j}$ is always set to $s_{i,j}^{\text{off}}$ at $\tau_{i,j}$ arrival time. These offline speeds $s_{i,j}^{\text{off}}$ are determined *before* the system execution and remain always constant at run-time. They may be simply set to the maximal processors speed s_{\max} , or they

can be determined by an *offline* energy-aware strategy, such that the one proposed in [26] for instance. These offline speeds *must ensure that all the deadlines are met* when the set of tasks is scheduled upon the m processors, even if every job of every task presents its WCET. Notice that, since each task generates an infinity of jobs, the method proposed in [26] determines a common speed for every task and assumes that every job $\tau_{i,j}$ inherits from the offline speed of τ_i at run-time.

MORA is based on reducing *online* (i.e., while the system is running) the execution speed $s_{i,j}$ of the jobs in order to provide energy savings *while still meeting all the deadlines*. To achieve this goal, MORA detects whenever the speed $s_{i,j}$ of an active job $\tau_{i,j}$ can safely be reduced by performing comparison between the schedule which is actually produced (called the *actual schedule* hereafter) and the *offline schedule* defined below. We will see in the remainder of this section that our algorithm MORA *always refers to this offline schedule* in order to produce the actual one.

Definition 1 (The offline schedule) *The offline schedule is the schedule produced by the considered scheduling algorithm on which every job of every task τ_i runs at its offline speed $s_{i,j}^{\text{off}}$ and presents its WCET.*

Figure 1.(a) depicts an example of an offline schedule and illustrates the notations that will be used throughout the paper. In this picture, a 5-tasks system is executed upon 2 processors, where only the first job of each task is represented. The characteristics of the tasks are the following (remember that $\tau_i = (C_i, D_i, T_i)$): $\tau_1 = (6, 14, 30)$, $\tau_2 = (6, 15, 35)$, $\tau_3 = (8, 16, 40)$, $\tau_4 = (2, 17, 45)$ and $\tau_5 = (6, 18, 50)$. Assuming Global-EDF, we have the following priority order: $\tau_{1,1} > \tau_{2,1} > \tau_{3,1} > \tau_{4,1} > \tau_{5,1}$. Furthermore, we assume in this example that the offline speed $s_{i,j}^{\text{off}}$ of every job $\tau_{i,j}$ is the maximal processors speed $s_{\max} = 1$.

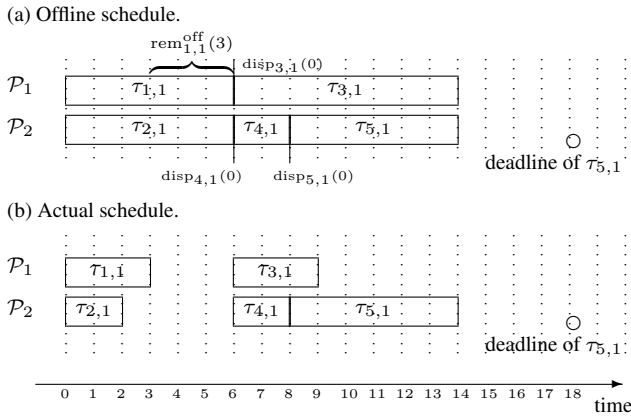


Figure 1. Offline and actual schedules.

At run-time, whenever any job is dispatched to any processor \mathcal{P}_ℓ in the offline schedule, MORA also dispatches

it to \mathcal{P}_ℓ in the actual one. That is, assuming the same set of tasks as in Figure 1.(a), Figure 1.(b) depicts the actual schedule that is produced if the *actual execution time* of the jobs $\tau_{1,1}, \dots, \tau_{5,1}$ are respectively 3, 2, 3, 2, 6. At any time t , we denote by $\text{rem}_{i,j}(t)$ and $\text{rem}_{i,j}^{\text{off}}(t)$ the worst-case *remaining* execution time of job $\tau_{i,j}$ at speed s_{\max} in the actual and offline schedule, respectively. We assume that these quantities are updated at run-time for every active job $\tau_{i,j}$. For instance in Figure 1 at time $t = 3$, we have $\text{rem}_{1,1}(3) = 0$ (since $\tau_{1,1}$ completes at time $t = 3$ in the actual schedule) and $\text{rem}_{1,1}^{\text{off}}(3) = 3$. Notice that from our definition of a processor speed, the worst-case *remaining* execution time of job $\tau_{i,j}$ at speed s in the actual and offline schedule is $\frac{\text{rem}_{i,j}(t)}{s}$ and $\frac{\text{rem}_{i,j}^{\text{off}}(t)}{s}$, respectively. We denote by $\text{disp}_{i,j}(t)$ the earliest time at which $\tau_{i,j}$ is dispatched in the offline schedule, when only the set of active jobs at time t in the offline schedule are considered. For instance in Figure 1.(a) we have $\text{disp}_{4,1}(0) = 6$ and $\text{disp}_{5,1}(0) = 8$. Finally, $\text{nextdisp}(\mathcal{P}_\ell, t)$ denotes the earliest instant after time t at which a job which is not completed in the actual schedule at time t is dispatched to \mathcal{P}_ℓ in the offline schedule. Again, only the set of active jobs at time t in the offline schedule are considered to compute $\text{nextdisp}(\mathcal{P}_\ell, t)$. For instance in Figure 1.(a) we have $\text{nextdisp}(\mathcal{P}_2, 2) = 6$ and $\text{nextdisp}(\mathcal{P}_1, 3) = 6$.

3.2. The α -queue

Since the jobs arrival times are unknown while considering the sporadic task model, computing and storing the entire offline schedule cannot be done *before the system execution*. Hence, our algorithm only stores and updates at run-time a *sufficient part* of the offline schedule. This kind of approach (i.e., using a dynamic data structure for embodying a sufficient part of the offline schedule) was previously proposed in [5]. As in [5], we call this data structure α -queue. The α -queue is a list that contains, at any time t , the worst-case remaining execution time $\text{rem}_{i,j}^{\text{off}}(t)$ of every active jobs $\tau_{i,j}$ in the offline schedule. This list is managed according to the following rules, which are widely inspired from [5].

α -Rule 1 *At any time, the α -queue is sorted by decreasing order of the job priorities, with the m highest priority jobs at the head of the queue.*

α -Rule 2 *Initially the α -queue is empty.*

α -Rule 3 *Upon arrival of a job $\tau_{i,j}$ at time t , $\tau_{i,j}$ inserts its WCET C_i into the α -queue in the correct priority position. This happens only once for each arrival, no re-insertion at return from preemptions.*

α -Rule 4 As time elapses, the m fields $\text{rem}_{i,j}^{\text{off}}(t)$ (if any) at the head of the α -queue are decreased with a rate proportional to the offline speeds $s_{i,j}^{\text{off}}$. Whenever one field reaches zero, that element is removed and the update continues, still with the m first elements (if any). Obviously, no update is performed when the α -queue is empty.

For the same reasons than those explained in [5], the following observation holds.

Observation 1 At any time t , the α -queue updated according to α -Rules 2–4 contains only the jobs that would be active at time t in the offline schedule. Moreover, the $\text{rem}_{i,j}^{\text{off}}(t)$ fields contain the worst-case remaining execution time of every active job $\tau_{i,j}$ at time t in the offline schedule.

By consulting the α -queue at any time t , MORA is able to get the required information about any active jobs $\tau_{i,j}$ in the offline schedule, i.e., its worst-case remaining execution time $\text{rem}_{i,j}^{\text{off}}(t)$, its next dispatching time $\text{disp}_{i,j}(t)$ and the next job dispatching time $\text{nextdisp}(\mathcal{P}_\ell, t)$ on any processor \mathcal{P}_ℓ . Due to the space limitation, we omitted the implementation details about the procedures which compute $\text{disp}_{i,j}(t)$ and $\text{nextdisp}(\mathcal{P}_\ell, t)$.

Notice that, as explained in [5], the dynamic reduction of $\text{rem}_{i,j}^{\text{off}}(t)$ from α -Rule 4 does not need to be performed at every clock cycle. Instead, for efficiency, we perform the reduction only *before* MORA modifies a speed, by taking into account the time elapsed since the last update. Formally, if Δt time units elapsed, the m fields at the head of the α -queue are updated as follows: $\text{rem}_{i,j}^{\text{off}}(t + \Delta t) \leftarrow \text{rem}_{i,j}^{\text{off}}(t) - s_{i,j}^{\text{off}} \cdot \Delta t$. The above approach relies on two facts: as we will see in the next section, the speed adjustment decisions will be taken only at job arrival time (i.e., the execution speed of the arriving job is set to its offline speed), job dispatching time in the offline schedule and whenever a processor is about to get idle in the actual schedule. Hence, it is necessary to have an accurate α -queue only at these instants. Second, between these instants, each task is effectively executed non-preemptively in the actual schedule.

3.3. Principle of MORA

As explained in Section 3.1, whenever a job is dispatched in the offline schedule, it is also dispatched in the actual one. However, as we will see below, MORA profits from an early job completion by starting the execution of some other jobs *earlier in the actual schedule than in the offline one*. As a result, when a job (say $\tau_{k,\ell}$) is dispatched at time t in the offline schedule (and thus also in the actual one), its worst-case remaining execution time $\text{rem}_{k,\ell}^{\text{off}}(t)$ could be lower than $\text{rem}_{k,\ell}^{\text{off}}(t)$ if it was executed earlier in the actual schedule. For example, Figure 2 depicts the same set of tasks than in Figure 1. At time $t = 2$, $\tau_{2,1}$ completes

in the actual schedule on processor \mathcal{P}_2 and leaves 4 unused time units. These 4 time units are reclaimed by starting the execution of $\tau_{5,1}$ (we will see below how MORA selects the job which profits from the slack time) and therefore, when $\tau_{5,1}$ is dispatched to \mathcal{P}_2 in the offline schedule at time $t = 8$, it is also dispatched to \mathcal{P}_2 in the actual one and we have $\text{rem}_{5,1}(8) < \text{rem}_{5,1}^{\text{off}}(8)$. The difference between these remaining execution times is called the *earliness* of the job and we denote it by $\epsilon_{k,\ell}(t) \stackrel{\text{def}}{=} \text{rem}_{k,\ell}^{\text{off}}(t) - \text{rem}_{k,\ell}(t)$. According to this earliness, whenever any job $\tau_{k,\ell}$ is dispatched in both schedules, its execution speed $s_{k,\ell}$ may safely be reduced to $s'_{k,\ell}$ so that $\frac{\text{rem}_{k,\ell}}{s'_{k,\ell}} = \frac{\text{rem}_{k,\ell} + \epsilon_{i,j}(t)}{s_{k,\ell}^{\text{off}}}$. Indeed, under this speed $s'_{k,\ell}$, $\tau_{k,\ell}$ would complete simultaneously in both schedules if it presents its WCET. This leads to the first rule of MORA.

Rule 1 Any job $\tau_{i,j}$ which is dispatched to any processor \mathcal{P}_ℓ at time t in the offline schedule is also dispatched to \mathcal{P}_ℓ at time t in the actual one and its execution speed $s_{i,j}$ is modified according to

$$s_{i,j} \leftarrow \hat{S} \left(\frac{\text{rem}_{i,j}(t) \cdot s_{i,j}^{\text{off}}}{\text{rem}_{i,j}^{\text{off}}(t)} \right) \quad (1)$$

The main idea of MORA can be summarized as follows. When any job completes in the actual schedule without consuming its WCET, the unused time may be reclaimed by starting the execution of any waiting job earlier; and since this waiting job receives additional time for its execution, it can thereby reduce its execution speed. Using this concept, Figure 2 depicts an example of how MORA takes advantage from an early job completion. When $\tau_{2,1}$ completes at time $t = 2$ in the actual schedule, MORA selects a waiting job (here, $\tau_{5,1}$) and executes it during the 4 time units left by $\tau_{2,1}$. Since $\tau_{5,1}$ is granted to use 4 additional time units, MORA reduces its execution speed $s_{5,1}$ so that its worst-case remaining execution time increases by 4 time units. The selected job is the one for which the resulting speed reduction leads to the highest energy saving. Formally, MORA selects a waiting job and decreases its execution speed as described by Rule 2.

(a) Offline schedule.

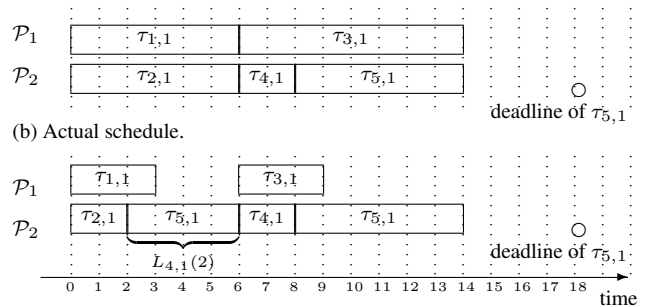


Figure 2. Rules 1 and 2 of MORA.

Rule 2 Whenever any processor \mathcal{P}_r is about to get idle at time t in the actual schedule,

Step 1. Use the α -queue to compute the next dispatching time $\text{nextdisp}(\mathcal{P}_r, t)$ on processor \mathcal{P}_r and proceed to the steps 2–5 for every waiting job $\tau_{i,j}$ at time t in the actual schedule.

Step 2. Compute the amount $L_{i,j}(t)$ of additional time units that $\tau_{i,j}$ could reclaim in the actual schedule if it was dispatched at time t , i.e.,

$$L_{i,j}(t) \stackrel{\text{def}}{=} \min(\text{nextdisp}(\mathcal{P}_r, t), \text{disp}_{i,j}(t)) - t$$

In Figure 2 we have $\text{nextdisp}(\mathcal{P}_2, 2) = 6$ and $L_{i,j}(2) = 4 \forall \tau_{i,j}$.

Step 3. Compute what would be the resulting execution speed $s'_{i,j}$ if $\tau_{i,j}$ was granted to use both its earliness and these $L_{i,j}$ additional time units, i.e., $s'_{i,j}$ is computed so that $\frac{\text{rem}_{i,j}(t)}{s'_{i,j}} = \frac{\text{rem}_{i,j}(t) + \epsilon_{i,j}(t)}{s_{i,j}^{\text{off}}} + L_{i,j}$, thus leading to

$$s'_{i,j} \leftarrow \widehat{S} \left(\frac{\text{rem}_{i,j}(t) \cdot s_{i,j}^{\text{off}}}{\text{rem}_{i,j}^{\text{off}}(t) + L_{i,j}(t) \cdot s_{i,j}^{\text{off}}} \right)$$

Step 4. Estimate what would be the resulting execution speed $s''_{i,j}$ if $\tau_{i,j}$ was not granted to use these $L_{i,j}(t)$ additional time units. According to Rule 1, $s_{i,j}$ will be modified to $s''_{i,j}$ when $\tau_{i,j}$ will be dispatched in the offline schedule (say at time t''). By assuming that $\tau_{i,j}$ will not be executed in the actual schedule until time t'' , we will have $\text{rem}_{i,j}(t'') = \text{rem}_{i,j}(t)$ and from Expression 1

$$s''_{i,j} \leftarrow \widehat{S} \left(\frac{\text{rem}_{i,j}(t) \cdot s_{i,j}^{\text{off}}}{\text{rem}_{i,j}^{\text{off}}(t)} \right)$$

Step 5. Compute the energy saving $\Delta E_{i,j}$ between execution at speed $s''_{i,j}$ and at speed $s'_{i,j}$:

$$\Delta E_{i,j} \leftarrow E_i \left(\frac{\text{rem}_{i,j}(t)}{s''_{i,j}}, s''_{i,j} \right) - E_i \left(\frac{\text{rem}_{i,j}(t)}{s'_{i,j}}, s'_{i,j} \right)$$

Step 6. Dispatch the job $\tau_{k,\ell}$ with the largest $\Delta E_{k,\ell}$ to processor \mathcal{P}_r . If $\Delta E_{i,j} \leq 0$ for all the waiting jobs, then dispatch the waiting job $\tau_{k,\ell}$ (if any) with the highest priority in order to complete it earlier and to potentially increase the length of future slack time.

Step 7. If there is a selected job $\tau_{k,\ell}$, set its execution speed $s_{k,\ell}$ to the computed one $s'_{k,\ell}$. Otherwise, turn the processor \mathcal{P}_r into the idle mode.

Notice that, if a processor is about to be idle in the actual schedule exactly when a job is dispatched in the offline one, only Rule 1 is applied. Algorithm 1 presents the pseudocode of MORA and we demonstrate its correctness in the following section.

Algorithm 1: MORA

```

1 Determine the offline speed  $s_{i,j}^{\text{off}}$  of every job  $\tau_{i,j}$  ;
2  $\alpha\text{-queue} \leftarrow \phi$  ;

At job arrival (say  $\tau_{i,j}$ ) at time  $t$ :
3 Update the  $\alpha$ -queue according to  $\alpha$ -Rule 4 ;
4 Insert the value of  $C_i$  into the  $\alpha$ -queue according to  $\alpha$ -Rule 3 ;
5 Set  $s_{i,j}$  to  $s_{i,j}^{\text{off}}$  ;

Whenever any processor  $\mathcal{P}_r$  is about to get idle at time  $t$ :
6 Update the  $\alpha$ -queue according to  $\alpha$ -Rule 4 ;
7 apply Rule 2 ;

Whenever any job  $\tau_{i,j}$  is dispatched to any processor  $\mathcal{P}_r$  in the offline schedule at time  $t$ :
8 Update the  $\alpha$ -queue according to  $\alpha$ -Rule 4 ;
9 if (a job  $\tau_{k,\ell} \neq \tau_{i,j}$  is running on  $\mathcal{P}_r$ ) then Preempt  $\tau_{k,\ell}$  ;
10 apply Rule 1 ;

```

3.4. Correctness of MORA

In this section, we formally prove that using MORA does not jeopardize the system schedulability.

Lemma 1 Let S be any preemptive and FJP global scheduling algorithm and let τ be any set of real-time tasks. Suppose that τ is scheduled by S while using MORA, and at time t during the system execution we have $\forall \tau_{i,j}$ and $\forall 0 \leq t' \leq t$:

$$\text{rem}_{i,j}(t') \leq \text{rem}_{i,j}^{\text{off}}(t')$$

Then, $\nexists t'$ with $0 \leq t' \leq t$ such that $\exists \tau_{i,j}$ running at time t' in the offline schedule and waiting at time t' in the actual one.

Proof The proof is obtained by contradiction. Suppose that at any time t' such that $0 \leq t' \leq t$, $\exists \tau_{i,j}$ running in the offline schedule and waiting in the actual one. It implies that at time t' in the offline schedule, there are at most $(m-1)$ jobs with an higher priority than $\tau_{i,j}$, whereas there are at least m such jobs in the actual one. In other words, there is at least one job (say $\tau_{k,\ell}$) at time t' with an higher priority than $\tau_{i,j}$, such that $\tau_{k,\ell}$ is completed in the offline schedule, and not in the actual one. For this job, it holds that $\text{rem}_{k,\ell}(t') > \text{rem}_{k,\ell}^{\text{off}}(t')$, leading to contradiction with our hypothesis. The property follows. ■

Lemma 2 Let S be any preemptive and FJP global scheduling algorithm and let τ be any set of real-time tasks. Suppose that τ is scheduled by S while using MORA, and at time t during the system execution we have $\forall \tau_{i,j}$ and $\forall 0 \leq t' \leq t$:

$$\text{rem}_{i,j}(t') \leq \text{rem}_{i,j}^{\text{off}}(t')$$

Then, $\nexists t'$ with $0 \leq t' \leq t$ such that $\exists \tau_{i,j}$ running at time t' in the offline schedule and such that the last speed modification of $\tau_{i,j}$ was performed according to Rule 2.

Proof The proof is obtained by contradiction. Suppose that at time t' with $0 \leq t' \leq t$, $\exists \tau_{i,j}$ running at time t' in the offline schedule and such that the last modification of $s_{i,j}$ was performed according to Rule 2. Let t_{actual} and t_{off} be the largest instants before time t' at which $\tau_{i,j}$ was dispatched in the actual and offline schedule, respectively. Notice that the case where $\tau_{i,j}$ is not dispatched before time t' in the actual schedule leads to a contradiction of Lemma 1. Therefore, only two cases may arise: (i) $t_{\text{actual}} \leq t_{\text{off}}$, in this case $s_{i,j}$ would have been modified at time t_{off} according to Rule 1, leading to a contradiction of our hypothesis, or (ii) $t_{\text{actual}} > t_{\text{off}}$, which leads to a contradiction of Lemma 1. The property follows. ■

Theorem 1 Let S be any preemptive and FJP global scheduling algorithm and let τ be any set of real-time tasks which is schedulable by S when every job $\tau_{i,j}$ is executed at its offline speed $s_{i,j}^{\text{off}}$. Then, every job deadline is still met when the system is scheduled by S while using MORA.

Proof The proof consists in showing that $\forall \tau_{i,j}$ we have

$$\text{rem}_{i,j}(d_{i,j}) \leq \text{rem}_{i,j}^{\text{off}}(d_{i,j}) \quad (2)$$

while using MORA. Indeed, since the offline schedule meets all the deadlines, we have $\text{rem}_{i,j}^{\text{off}}(d_{i,j}) = 0 \forall \tau_{i,j}$. Therefore, having $\text{rem}_{i,j}(d_{i,j}) \leq \text{rem}_{i,j}^{\text{off}}(d_{i,j})$ leads to $\text{rem}_{i,j}(d_{i,j}) = 0 \forall \tau_{i,j}$, meaning that the actual schedule also meets all the deadlines.

Initially at time $t = 0$, we obviously have $\text{rem}_{i,j}(0) = \text{rem}_{i,j}^{\text{off}}(0) \forall \tau_{i,j}$. Now, let $t > 0$ be any instant and suppose that $\forall \tau_{i,j}$ and $\forall 0 \leq t' \leq t$ we have $\text{rem}_{i,j}(t') \leq \text{rem}_{i,j}^{\text{off}}(t')$. We prove in the following that it yields

$$\text{rem}_{i,j}(\text{next}(t)) \leq \text{rem}_{i,j}^{\text{off}}(\text{next}(t)) \quad \forall \tau_{i,j} \quad (3)$$

where $\text{next}(t)$ denotes the earliest instant after time t such that one of the following events occurs: arrival of a job, deadline of a job, completion of a job in the actual schedule or in the offline schedule, dispatching of a job in the actual schedule or in the offline schedule. Obviously if Inequality 3 holds then Inequality 2 also holds since $\text{next}(t)$ can denote every job deadline.

From the definition of $\text{next}(t)$, every processor of both schedules is either idle or it executes one and only one job during any time interval $[t, \text{next}(t)]$. In other words, the state (waiting or running) of any active jobs in any schedule does not change during any time interval $[t, \text{next}(t)]$. As a result, the following relations hold at time t :

- For any waiting job $\tau_{i,j}$ in the actual schedule:

$$\text{rem}_{i,j}(\text{next}(t)) = \text{rem}_{i,j}(t) \quad (4)$$

- For any waiting job $\tau_{i,j}$ in the offline schedule:

$$\text{rem}_{i,j}^{\text{off}}(\text{next}(t)) = \text{rem}_{i,j}^{\text{off}}(t) \quad (5)$$

- For any running job $\tau_{i,j}$ in the actual schedule:

$$\text{rem}_{i,j}(\text{next}(t)) \leq \text{rem}_{i,j}(t) \quad (6)$$

The first part of the proof shows that Inequality 3 holds for every waiting job at time t in the actual schedule and the second part shows that it also holds for every running job at time t in the actual schedule.

Part 1. Let $\tau_{k,\ell}$ be any waiting job at time t in the actual schedule. From Lemma 1, we know that $\tau_{k,\ell}$ is also waiting at time t in the offline one and since by hypothesis $\text{rem}_{k,\ell}(t) \leq \text{rem}_{k,\ell}^{\text{off}}(t)$, we know from Equalities 4 and 5 that $\text{rem}_{k,\ell}(\text{next}(t)) \leq \text{rem}_{k,\ell}^{\text{off}}(\text{next}(t))$. The property follows.

Part 2. Let $\tau_{k,\ell}$ be any running job at time t in the actual schedule. Regarding its execution speed $s_{k,\ell}$, only two cases may occur: its last modification was performed by Rule 1 (case 1) or by Rule 2 (case 2).

Case 1. $\tau_{k,\ell}$ is running at time t in the actual schedule and the last modification of $s_{k,\ell}$ was performed according to Rule 1 when it was dispatched in the offline schedule (say at time $t_{\text{off}} \leq t$). By hypothesis, we have $\text{rem}_{k,\ell}(t_{\text{off}}) \leq \text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})$ and we know that $\tau_{k,\ell}$ is executed non-preemptively in both schedules during the time interval $[t_{\text{off}}, t]$. Indeed, if it was preempted in the offline schedule, it would have been also preempted in the actual one according to Rule 1. However in the actual schedule, $\tau_{k,\ell}$ is running at times t_{off} and t . Therefore, its speed would have been modified according to Rule 2 at its re-dispatching time if it was preempted during $[t_{\text{off}}, t]$. As a result, from our interpretation of the processor speed we get

$$\text{rem}_{k,\ell}(\text{next}(t)) = \text{rem}_{k,\ell}(t_{\text{off}}) - s_{k,\ell} \cdot (\text{next}(t) - t_{\text{off}}) \quad (7)$$

and

$$\text{rem}_{k,\ell}^{\text{off}}(\text{next}(t)) = \text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}}) - s_{k,\ell}^{\text{off}} \cdot (\text{next}(t) - t_{\text{off}}) \quad (8)$$

After the speed modification by Rule 1 at time t_{off} , we know from Expression 1 that $s_{k,\ell} = \frac{\text{rem}_{k,\ell}(t_{\text{off}})}{\text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})} \cdot s_{k,\ell}^{\text{off}}$ and Equality 7 can be rewritten as

$$\text{rem}_{k,\ell}(\text{next}(t)) = \text{rem}_{k,\ell}(t_{\text{off}}) - \frac{\text{rem}_{k,\ell}(t_{\text{off}})}{\text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})} \cdot s_{k,\ell}^{\text{off}} \cdot (\text{next}(t) - t_{\text{off}})$$

Finally, notice that multiplying the right-hand side of the above Equality by $\frac{\text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})}{\text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})}$ leads to the right-hand side of Equality 8. Since by hypothesis $\text{rem}_{k,\ell}(t_{\text{off}}) \leq \text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})$, we have $\frac{\text{rem}_{k,\ell}(t_{\text{off}})}{\text{rem}_{k,\ell}^{\text{off}}(t_{\text{off}})} \geq 1$ and therefore $\text{rem}_{k,\ell}(\text{next}(t)) \leq \text{rem}_{k,\ell}^{\text{off}}(\text{next}(t))$. The property follows.

Case 2. $\tau_{k,\ell}$ is running at time t in the actual schedule and the last modification of $s_{k,\ell}$ was performed according to Rule 2. Therefore, we know from Lemma 2 that $\tau_{k,\ell}$ is waiting at time t in the offline schedule, and since by hypothesis $\text{rem}_{k,\ell}(t) \leq \text{rem}_{k,\ell}^{\text{off}}(t)$, we know from Equalities 5 and 6 that $\text{rem}_{k,\ell}(\text{next}(t)) \leq \text{rem}_{k,\ell}^{\text{off}}(\text{next}(t))$. The theorem follows. ■

4. Simulation results

In this section, we compare the effectiveness of MORA with other energy-aware algorithms. However, it is meaningful to only compare MORA with approaches that consider the same models of computation and the most related paper to ours is [26], where two methods with the same task and platform model are proposed. However, these two methods do not take into account the application-specific parameter e_i of task τ_i . The first method proposed in [26] (that we denote by OFF hereafter) is an offline speed determination technique for Global-EDF which determines a unique and constant speed s^{off} for all the processors such that all the job deadlines are met under this speed. In our simulations, this OFF method is used by MORA in order to provide the offline speed $s_{i,j}^{\text{off}}$ of every job $\tau_{i,j}$, i.e., s^{off} is determined at line 1 of Algorithm 1 and $s_{i,j}^{\text{off}}$ is set to s^{off} between lines 4 and 5. The second method proposed in [26] is the MOTE algorithm. At run-time, it anticipates the coming idle instants in the schedule and adjusts the speed of the processors accordingly, i.e., it reduces the processors speed in order to minimize the proportion of time during which the system is idle. Since this algorithm is also based on the concept of the offline speeds, we consider that OFF is also used to provide it. Although MORA could also be compared with frame-based scheduling algorithms (since the sporadic task model is a generalization of the frame-based task model), we do not perform such comparisons in this paper.

In our simulations, we schedule *periodic implicit-deadline systems* (i.e., $\forall \tau_i$, T_i is here the *exact* inter-arrival delay between successive jobs and $D_i = T_i$). The energy consumption of each generated system is computed by simulating three methods: MOTE, MORA and MORAOTE, i.e., a combination of the MOTE and MORA. Indeed, since these algorithms do not interfere with each other, the MOTE rule can be applied *on the offline speeds* just *before applying Rule 1* of MORA (i.e., between lines 9 and 10 of Algorithm 1). Although the implementation details of MORAOTE are omitted here due to the space limitation, we will see in our simulation results that this combination always improves the provided energy savings. The consumptions provided by these three methods are compared with the consumption of the MAX method, where all the jobs are executed at the maximal processors speed $s_{\text{max}} = 1$. That is, we consider that the consumption by MAX is 100% and

the consumptions of the other methods are normalized.

In every simulation, we generated 100 set of tasks with a total density $\delta_{\text{sum}}(\tau)$ within $[d, d + 0.05]$ where $d = 0, 0.05, \dots, 9.95$, leading to an amount of 20000 generated task sets for each simulation. The upper bound on $\delta_{\text{sum}}(\tau)$ (i.e., 10) was chosen in order to cover a large number of systems while keeping the simulation time reasonable. For a given *total density*, tasks densities δ_i are uniformly generated within $[0.01, D_{\text{max}}]$ until the total density $\delta_{\text{sum}}(\tau)$ reaches the expected one (the upper bound D_{max} on the tasks density will be discussed later). Notice that the number n of tasks is not fixed beforehand, i.e., it depends on this step that generates task densities. Next, other task parameters C_i , D_i and T_i are randomly generated according to their respective density δ_i . Finally, the application-specific parameters e_i are uniformly chosen in $[0.8, 1.2]$ so that the consumption of the tasks varies between 80% and 120% of the power of the measured benchmark.

Once a set of tasks is generated, it is executed during 100 hyper-periods (i.e. the least common multiple of the task periods) by the four methods MAX, MOTE, MORA and MORAOTE. This upper bound on systems execution time was chosen to ensure that every task generates at least 100 jobs (for the same reason as those mentioned above). During each system execution, the actual execution time of every job $\tau_{i,j}$ is uniformly generated in $[\frac{C_i}{10}, C_i]$. This lower bound $\frac{C_i}{10}$ was chosen in order to reflect the fact that a job may take up to 10 times less than its WCET. Finally, for every generated task set τ , the number m of processors must be sufficient to schedule τ by MAX without missing any deadline. Hence, we set m to the lowest integer that passes one of the following EDF-schedulability tests: the *density-based* test [19], the *load-based* test [12] and the test denoted Test 13 in [8]. Simulations were performed while considering different scheduling algorithms (Global-EDF and Global-DM) and various processor models. However, due to the space limitation, we only depict in this paper the results provided by Global-EDF on Intel XScale processors (outlined in Table 1 page 2).

Observation 2 *The effectiveness of both MORA and MOTE mainly relies on the ratio $\frac{m}{n}$, but antagonistically.*

This observation stems from the fact that MORA saves energy via the waiting jobs whereas MOTE profits from the absence of waiting jobs. When $\frac{m}{n}$ tends to 1, jobs tend to never wait for a free processor and MOTE therefore provides significant energy savings whereas the effectiveness of MORA is almost null. On the other hand when $\frac{m}{n}$ tends to 0, processors tend to consecutively execute several distinct jobs and jobs are often waiting. As a result, MORA is often able to reclaim unused time and provides important energy savings whereas the effectiveness of MOTE is negligible.

According to our task generation process, we are not able to directly set the ratio $\frac{m}{n}$ to any given value. However, the number m of processors is obtained by using a combination of *sufficient* schedulability tests and the accuracy of these tests mainly relies on $\delta_{\max}(\tau)$. Basically, the ratio $\frac{m}{n}$ increases as $\delta_{\max}(\tau)$ becomes larger and since the generated task sets are more likely to have a large $\delta_{\max}(\tau)$ when the upper bound D_{\max} is high, we can indirectly control the ratio $\frac{m}{n}$ via D_{\max} . The Y-axis of Figure 3 represents the ratio $\frac{m}{n}$ obtained from the used schedulability tests when $\delta_{\text{sum}}(\tau)$ varies within $[0, 10]$ and D_{\max} varies within $[0.1, 1]$ with a step of 0.1.

For every D_{\max} multiple of 0.1 within $[0.1, 1]$, 20000 set of tasks were generated by the generation process described above and the resulting average consumptions of the MOTE, MORA and MORAOTE are depicted in Figure 4. The Y-axis is the average energy consumption of every method compared with the MAX method (in %) and the X-axis is the corresponding value of D_{\max} during the simulation. As we can see, Figures 3 and 4 clearly corroborate Observation 2. Moreover, Figure 4 shows that MORA can save up to 32% of energy (in average) over the MAX method (for $D_{\max} = 0.1$) and the algorithm MORAOTE provides important energy savings for various values of D_{\max} . Notice that a part of the energy savings is explained by the use of the OFF method, which leads MOTE and MORA to an energy savings of about 10% when $\frac{m}{n}$ tends to 0 and 1, respectively. Furthermore, although other processor models and scheduling algorithms led to different average consumptions, the *evolution* of the consumption with respect to D_{\max} remains similar than in Figure 4.

5. Conclusion

In this paper, we propose a slack reclamation scheme called MORA which reduces the energy consumption while scheduling a set of *sporadic constrained-deadline tasks* by a *global, preemptive* and FJP algorithm on a *fixed number of DVFS-identical processors*. According to [15] and to the best of our knowledge, we are the firsts to address such approach in this context. The proposed algorithm MORA exploits early job completions at run-time by starting the execution of the next waiting jobs at a lower speed. Compared with other reclaiming algorithms such that the DRA proposed in [5], MORA takes into account the application-specific consumption profile of the tasks in order to improve the energy saving that it provides. Moreover, we proved that using MORA does not jeopardize the system schedulability and we show in our simulations that it can save up to 32% of energy (in average) compared to execution without using any energy-aware algorithm.

In our future works, we aim to specialize MORA so that it will take into account more practical constraints such that

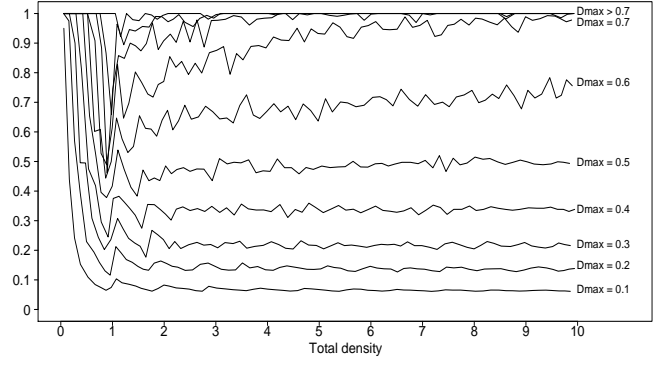


Figure 3. Ratio $\frac{m}{n}$ for different values of D_{\max}

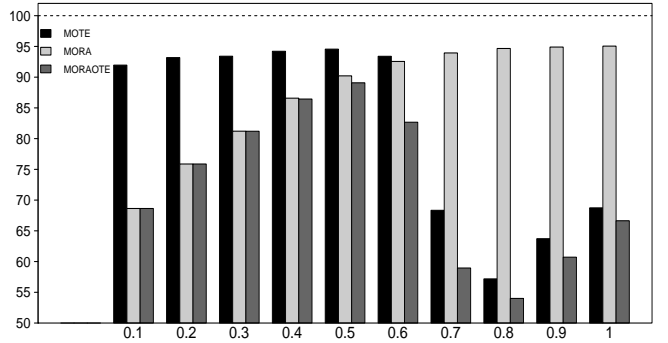


Figure 4. Average consumptions of MOTE, MORA and MORAOTE for various values of D_{\max} under Global-EDF on Intel XScale processors.

preemption costs, migration costs and time overheads due to the multiple frequency switching. Moreover, we aim to extend our processor model in order to handle the various idle and sleep modes of the processors and to take into account the energy costs due to frequency switching. In other future works, we also aim to propose a new multiprocessor reclamation scheme which anticipates the early completion of jobs for further reducing the CPU speed. This approach will be based on statistical informations about tasks that are assumed to be known *a priori*. Some *uniprocessor* energy-aware algorithms already exploit this concept (see the AGR algorithm proposed in [5] for instance).

References

- [1] Intel XScale Microarchitecture: Benchmarks, 2005. <http://web.archive.org/web/20050326232506/developer.intel.com/design/intelxscale/benchmarks.htm>.
- [2] T. A. AlEnawy and H. Aydin. Energy-aware task allocation for rate monotonic scheduling. In *Proceedings of the*

- 11th IEEE Real-time and Embedded Technology and Applications Symposium, pages 213–223, 2005.
- [3] J. Anderson and S. Baruah. Energy-efficient synthesis of EDF-scheduled multiprocessor real-time systems. *International Journal of Embedded Systems*, 4(1), 2008.
 - [4] H. Aydin, R. Melhem, D. Mossé, and P. Mejía-Alvarez. Determining optimal processor speeds for periodic real-time tasks with different power characteristics. In *Proceedings of the IEEE EuroMicro Conference on Real-Time Systems*, pages 225–232, 2001.
 - [5] H. Aydin, R. Melhem, D. Mossé, and P. Mejía-Alvarez. Power-aware scheduling for periodic real-time tasks. *IEEE Transactions on Computers*, 53(5):584–600, 2004.
 - [6] H. Aydin and Q. Yang. Energy-aware partitioning for multiprocessor real-time systems. In *Proceedings of 17th International Parallel and Distributed Processing Symposium*, pages 113–121, 2003.
 - [7] T. Baker. Multiprocessor EDF and deadline monotonic schedulability analysis. In *Proceedings of the 24th IEEE International Real-Time Systems Symposium*, pages 120–129, December 2003.
 - [8] T. Baker and S. Baruah. *Schedulability Analysis of Multiprocessor Sporadic Task Systems*. In Handbook of Real-Time and Embedded Systems, Sang H. Son, Insup Lee, and Joseph Y-T Leung (eds). Chapman Hall/ CRC Press, December 2006.
 - [9] N. Bansal, T. Kimbrel, and K. Pruhs. Dynamic speed scaling to manage energy and temperature. In *Proceedings of the Symposium on Foundations of Computer Science*, pages 520–529, 2004.
 - [10] S. Baruah and J. Anderson. Energy-aware implementation of hard-real-time systems upon multiprocessor platform. In *Proceedings of the 16th International Conference on Parallel and Distributed Computing Systems*, pages 430–435, August 2003.
 - [11] S. Baruah and J. Anderson. Energy-efficient synthesis of periodic task systems upon identical multiprocessor platforms. In *Proceedings of the 24th International Conference on Distributed Computing Systems*, pages 428–435, Tokyo, Japan, March 2004. IEEE Computer Society Press.
 - [12] S. Baruah and T. Baker. Schedulability analysis of global EDF. *Real Time Systems*, Accepted for publication, 2008.
 - [13] V. Bertin and J. Goossens. Multiprocessor global scheduling on frame-based DVFS systems. In I. Puaud, editor, *The 29th IEEE Real-Time Systems Symposium, WiP proceedings*, pages 21–24, 2008.
 - [14] J.-J. Chen, H.-R. Hsu, K.-H. Chuang, C.-L. Yang, A.-C. Pang, and T.-W. Kuo. Multiprocessor energy-efficient scheduling with task migration considerations. In *Proceedings of the 16th Euromicro Conference on Real-Time Systems (ECRTS'04)*, pages 101–108, 2004.
 - [15] J. J. Chen and C. Kuo. Energy-efficient scheduling for real-time systems on dynamic voltage scaling (DVS) platforms. In *the 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)*, pages 28–38, August 21-24 2007.
 - [16] J.-J. Chen and T.-W. Kuo. Multiprocessor energy-efficient scheduling for real-time tasks. In *International Conference on Parallel Processing*, pages 13–20, 2005.
 - [17] J.-J. Chen, C.-Y. Yang, and T.-W. Kuo. Slack reclamation for real-time task scheduling over dynamic voltage scaling multiprocessors. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, volume 1, pages 358–367, June 2006.
 - [18] J.-J. Chen, C.-Y. Yang, and T.-W. Kuo. Slack reclamation for real-time task scheduling over dynamic voltage scaling multiprocessors. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC)*, Taichung, Taiwan, 2006.
 - [19] J. Goossens, S. Funk, and S. Baruah. Priority-driven scheduling of periodic task systems on uniform multiprocessors. *Real Time Systems*, 25:187–205, 2003.
 - [20] <http://www.pdadb.net>.
 - [21] Intel. Intel PXA27x Processor Family, Design guide, May 2005.
 - [22] S. Irani, S. Shukla, and R. Gupta. Algorithms for power savings. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 37–46, 2003.
 - [23] T. Ishihara and H. Yasuura. Voltage scheduling problem for dynamically variable voltage processors. In *International Symposium on Low Power Electronics and Design*, pages 197–202, 1998.
 - [24] G. Magklis, G. Semeraro, D. H. Albonesi, S. G. Dropsho, S. Dwarkadas, and M. L. Scott. Dynamic frequency and voltage scaling for a multiple-clock-domain microprocessor. In *IEEE Micro*, volume 23, pages 62–68, 2003.
 - [25] P. Mejía-Alvarez, E. Levner, and D. Mossé. Adaptive scheduling server for power-aware real-time tasks. *ACM Transactions on Embedded Computing Systems*, 2(3):284–306, 2004.
 - [26] V. Nelis, J. Goossens, N. Navet, R. Devillers, and D. Milojevic. Power-aware real-time scheduling upon identical multiprocessor platforms. In *IEEE International Conference on Sensor Networks Ubiquitous and Trustworthy Computing*, pages 209–216, June 2008.
 - [27] P. Pillai and K. Shin. Real-time dynamic voltage scaling for low powered embedded systems. *Operating Systems Review*, 35:89–102, October 2001.
 - [28] Y. Shin and K. Choi. Power conscious fixed priority scheduling for hard real-time systems. In *Design Automation Conference*, pages 134–139, 1999.
 - [29] E. Talpes and D. Marculescu. Toward a multiple clock/voltage island design style for power-aware processors. In *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, volume 13, pages 591–603, 2005.
 - [30] R. Xu, R. Melhem, and D. Mosse. A unified practical approach to stochastic DVS scheduling. In *EMSOFT*, pages 37–46, 2007.
 - [31] C.-Y. Yang, J.-J. Chen, and T.-W. Kuo. An approximation algorithm for energy-efficient scheduling on a chip multiprocessor. In *Proceedings of the 8th Conference of Design, Automation, and Test*, pages 468–473, 2005.
 - [32] F. Yao, A. Demers, and S. Shenker. A scheduling model for reduced CPU energy. In *Proceedings of the 36th IEEE Annual Foundations of Computer Science*, pages 374–382, 1995.
 - [33] F. Zhang and S. Chanson. Processor voltage scheduling for real-time tasks with non-preemptible sections. In *23th Real-Time Systems Symposium*, pages 235–245, 2002.

- [34] D. Zhu, R. Melhem, and B. Childers. Scheduling with dynamic voltage/speed adjustment using slack reclamation in multi-processor real-time systems. In *Proceedings of IEEE 22th Real-Time System Symposium*, pages 84–94, 2001.
- [35] J. Zhuo and C. Chakrabarti. System-level energy-efficient dynamic task scheduling. In *ACM/IEEE Design Automation Conference*, pages 628–631, June 2005.